OHIO RIVER BASIN PRECIPITATION FREQUENCY STUDY

Update of *Technical Paper No. 40, NWS HYDRO-35* and *Technical Paper No. 49*

Eleventh Progress Report
1 April 2002 through 30 June 2002

Hydrometeorological Design Studies Center
Hydrology Laboratory

Office of Hydrologic Development
U.S. National Weather Service
Silver Spring, Maryland

July 2002

DISCLAIMER

TABLE OF CONTENTS

OHIO RIVER BASIN PRECIPITATION FREQUENCY STUDY

Update of *Technical Paper No. 40, NWS HYDRO-35* and *Technical Paper No. 49*

1. Introduction

The Hydrometeorological Design Studies Center (HDSC), Hydrology Laboratory, Office of Hydrologic Development, U.S. National Weather Service is updating its precipitation frequency estimates for the Ohio River Basin and surrounding states. Current precipitation frequency estimates for this area are contained in *Technical Paper No. 40* "Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years" (Hershfield 1961), *NWS HYDRO-35* "Five- to 60-minute precipitation frequency for the eastern and central United States" (Frederick et al 1977) and *Technical Paper No. 49* "Two- to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States" (Miller et al 1964). The new study includes collecting data and performing quality control, compiling and formatting datasets for analyses, selecting applicable frequency distributions and fitting techniques, analyzing data, mapping and preparing reports and other documentation.

The study will determine annual all-season precipitation frequencies for durations from 5 minutes to 60 days, for return periods from 2 to 1000 years. The study will review and process all appropriate rainfall data for the study area and use accepted statistical methods. The study results will be published as a Volume of NOAA Atlas 14 on the Internet with the additional ability to download digital files.

The study will produce estimates for 13 states. Parts of nine additional bordering states are included to ensure continuity across state borders. The Susquehanna River and Delaware River Basins are included in the study area. The core and border areas, as well as tentative regions used in the analysis, are shown in Figure 1.

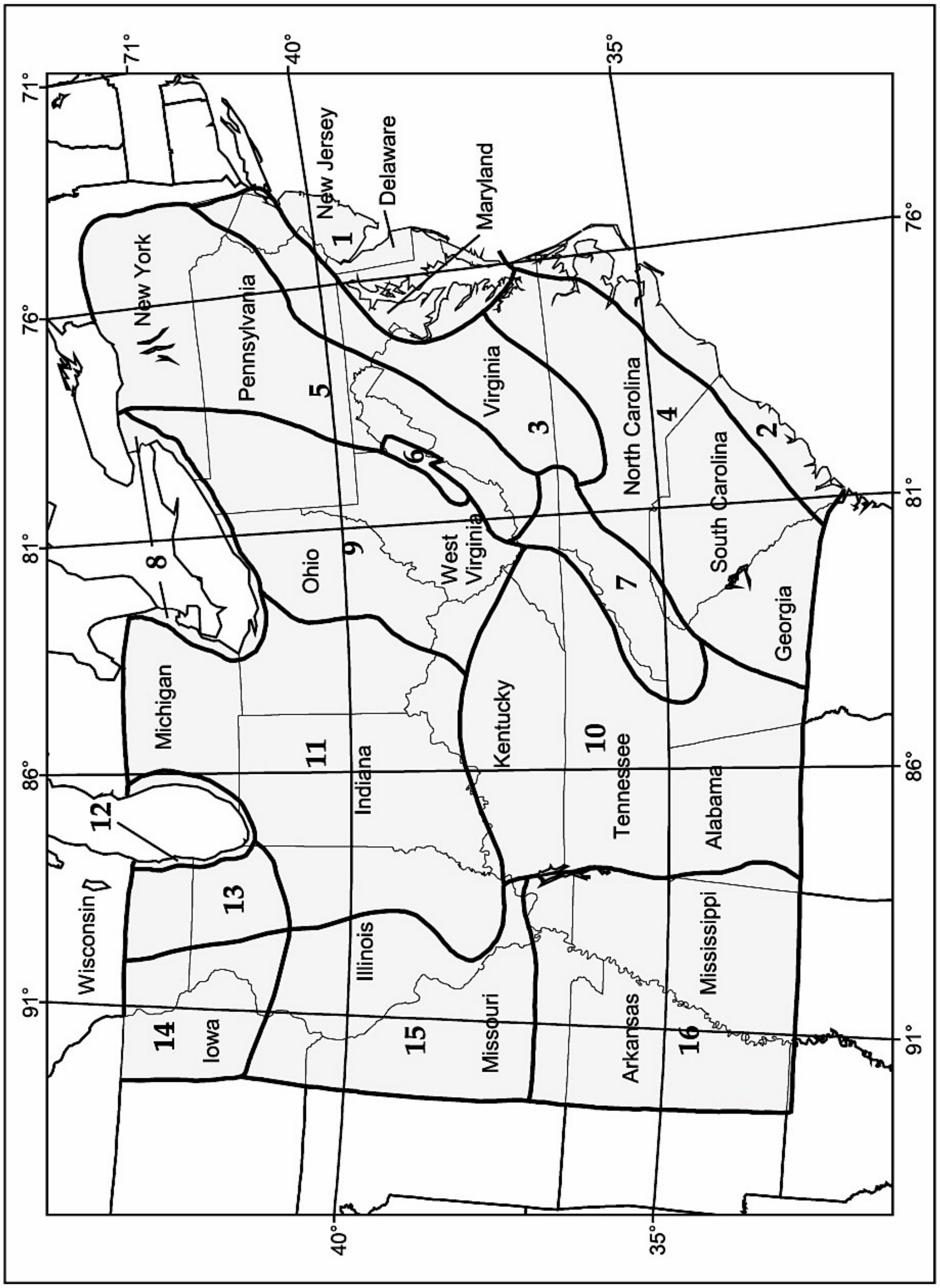


Figure 1. Ohio River Basin Precipitation Frequency study area and region boundaries.

2. Highlights

The Huntington, Louisville and Nashville U.S. Army Corps of Engineers (COE) districts had provided HDSC with data extending through December 2000. The data have been added and quality controlled. Additionally, after the initial L-moment results were computed, station discordancies were analyzed. Additional information is provided in Section 4.1, Data Collection and Quality Control.

Software has been developed to screen the 1-day annual maxima series for large gaps in time using specific “Gap Check” criteria. Stations have been adjusted where appropriate to produce more congruent data records. Additional information is provided in Section 4.2, Software Updates.

An initial run of the L-moment software has been completed for all daily and hourly durations and all regions. A region by region examination of heterogeneity is currently underway. Additional information is provided in Section 4.3, L-moment Analysis.

Twelve monthly percentage grid maps have been created for the temporal distribution analysis. Some generalized spatial patterns and monthly differences are evident in the percentage grid maps. Additional information is provided in Section 4.4, Temporal Distribution.

Several changes were made to the Precipitation Frequency Data Server (PFDS). Most important was the added function for selecting, either via a map or a list, an observing site. This function will allow for review of the point-precipitation frequency estimates before the interpolated grids are finalized. Additional information is provided in Section 4.5, Precipitation Frequency Data Server.

Depth Area Duration (DAD) reductions for areas from 10 to 400 square miles are being updated for the entire United States and will be presented in a separate volume of NOAA Atlas 14. The method to be used for computing the DAD curves has been selected. If additional dense-area-networks are available, they will be added to our database. Additional information is provided in Section 4.6, Spatial Relations (Depth Area Duration Study).

Decisions have been made to exclusively publish the study results electronically to avoid printing expenses, and to publish monthly patterns of extreme precipitation but not compute monthly frequency estimates. Additional information is provided in Section 5, Issues.

3. Status

3.1 Project Task List

The following checklist shows the components of each task and an estimate of the percent completed per task. Past status reports should also be referenced for additional information.

Ohio River Basin study checklist [estimated percent complete]:

Data Collection, Formatting and Quality Control [100%]:
- Multi-day
- Daily
- Hourly
- 15-minute
- N-minute

The Huntington, Louisville and Nashville U.S. Army Corps of Engineers (COE) districts provided data through December 2000. The appended data have been quality controlled. Additionally, after the initial L-moment results were computed, station discordancies were analyzed. As data issues may arise in subsequent tasks, quality control is essentially a continuous process.

L-Moment Analysis/Frequency Distribution for 5 minute to 60 days and 2 to 1000 years [20%]:
- Multi-daily
- Daily
- Hourly
- 15-minute
- N-minute

Software was developed to screen the 1-day annual maxima series for large gaps using specific “Gap Check” criteria. Stations have been adjusted where appropriate to produce more congruent data records. An initial run of the L-moment software has been completed for all daily and hourly durations in all regions.

Spatial Interpolation [0%]
- Create grids of interpolated means for each duration using PRISM

Peer Reviews [0%]:
- External peer review of point precipitation frequency estimates
- External peer review of spatial interpolation grids

Precipitation Frequency Maps [0%]
- Create smoothed regional growth factor (RGF) grids using GRASS
- Multiply appropriate RGF and distributed mean grids to produce precipitation frequency grids for the durations of 5-min, 10-min, 15-min, 30-min, 1-hr, 2-hr, 3-hr, 6-hr, 12-hr, 24-hr, 48-hr, 4-day, 7-day, 10-day, 20-day, 30-day, 45-day and 60-day at the 2-yr, 5-yr, 10-yr, 25-yr, 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr return frequencies for a total of 162 maps
- Apply study-wide conversion factor to the 1-hour precipitation frequency grids to calculate the n-minute (5-, 10-, 15-, and 30-minute) grids
- Perform internal consistency checks (comparing rasters of sequential duration and frequency)

Data Trend Analysis [0%]
- Analyze linear trends in annual maxima and variance over time
- Analyze shift in means of annual maxima between two time periods (i.e., test the equality of 2 population distribution means)

Temporal Distributions of Extreme Rainfall [20%]
- Assemble hourly data by quartile of greatest precipitation amount and convert to cumulative rainfall amounts for each region
- Sort, average and plot time distributions of hourly maximum and median events by storm area, quartile and duration

Twelve monthly percentage grid maps have been created for the temporal distribution analysis. Some generalized spatial patterns and monthly differences are evident in the percentage grid maps.

Deliverables [20%]
- Prepare data and documentation for web delivery

Several changes were made to the Precipitation Frequency Data Server (PFDS). Most important was the added function for selecting an observing site. This allows review of the point-precipitation frequency estimates before the interpolated grids are finalized.

Additional Work:

Spatial Relations (Depth Area Duration Study) [25%]
- Obtain hourly data from dense-area reporting networks
- QC and format data from dense networks
- Compute maximum and average annual areal depth for each duration from stations from each network
- Compute maximum to average depth ratio for all durations and networks and plot
- Prepare curves of best fit (depth area curves) for each duration and network

Depth Area Duration (DAD) reductions for areas from 10 to 400 square miles are being updated for the entire United States and will be presented in separate volume of NOAA Atlas 14.

4. Progress in this Reporting Period

4.1 Data Collection and Quality Control

The Huntington, Louisville and Nashville U.S. Army Corps of Engineers (COE) districts had provided data extending through December 2000. The data have been added and quality controlled.

The preliminary L-moment analysis results indicated that some stations had high discordancy values. Basically, these discordant stations represented statistical outliers. Discordant stations were analyzed, and the results of this analysis indicated that regionalization issues or sampling error issues may be driving the discordancy. The data check for the discordant stations did not reveal any data problems.

4.2 Software Updates

Some stations included in the Study had multiple missing years. Large gaps (i.e., sequential missing years) in an annual maxima series cause concern about the data series consistency. Using newly developed software, we screened all data records for large gaps using specific “Gap Check” criteria before data were used for L-moment analysis. The criteria are:

1. Size criteria: Software will delete a beginning short data segment and gap segment if the data years of the short segment and the gap size meet one of the following conditions (and only if it also meets the subsequent criteria as well):

data years size	gap size
1 yr	10 yrs
2 yrs	15 yrs
3 yrs	18 yrs
4 yrs	20 yrs
5 - 10 yrs	>30 yrs
10 - 15 yrs	>40 yrs
15 yrs	>50 yrs

2. High value criteria: If one of the highest annual maximum values for that station occurs in the short data segment, the software will not delete that segment but copy that station’s annual maximum series to a log file that will be manually checked.

3. 20 year criteria: The software will not delete data segments if it leaves less than 20 years of data for that station, but will copy that station's annual maximum series to the log file.

4. Middle or end gap criteria: Stations with large gaps in the middle or towards the end of their record will be copied to the log file. Large gaps are defined as 10 years or more.

5. Sporadic gap criteria: If there are two or more sequential gaps of 5 or more missing years separated by 5 or less years of data at a station, then the annual maximum series will be copied to the log file.

Station records with gaps were flagged by the software and examined on a case by case basis using an approach which preserved as much of the data as possible. Nearby stations were inspected for concurrent data years to fill in the gap if they passed the statistical test for consistency. Latitude, longitude and elevation were taken into account when examining nearby stations. Also, if there were sufficient years in each data segment, a t-test was conducted on the two segments to assess the statistical integrity of the data record. To produce more congruent data records for analysis, station record length was occasionally adjusted.

4.3 L-moment Analysis

An initial run of the L-moment software has been completed for all daily and hourly durations for all regions. The preliminary results indicate that most regions are statistically homogeneous. A homogenous region has identical frequency distributions, apart from the site-specific index flood factor, for every site. Table 1 shows the results for the hourly data.

The four regions which are heterogeneous in the 24-hour/1-day durations are being evaluated; regional subdivision may occur where appropriate. It is expected that heterogeneity issues in shorter (i.e., 2-hour) as well as longer durations will be eliminated once the regions are examined according to the statistical analysis of the 24-hour/1-day results. Factors defining the regions include 1) the season of highest precipitation, 2) the type of precipitation (e.g., general storm, convective, tropical storms or hurricanes), 3) the climate, 4) the topography and 5) the statistical homogeneity of these factors. Essentially, these are regions that have similar 24-hour/1-day extreme precipitation characteristics.

Table 1. Results from hourly L-moment Analysis (values >= 2.00 are heterogeneous).

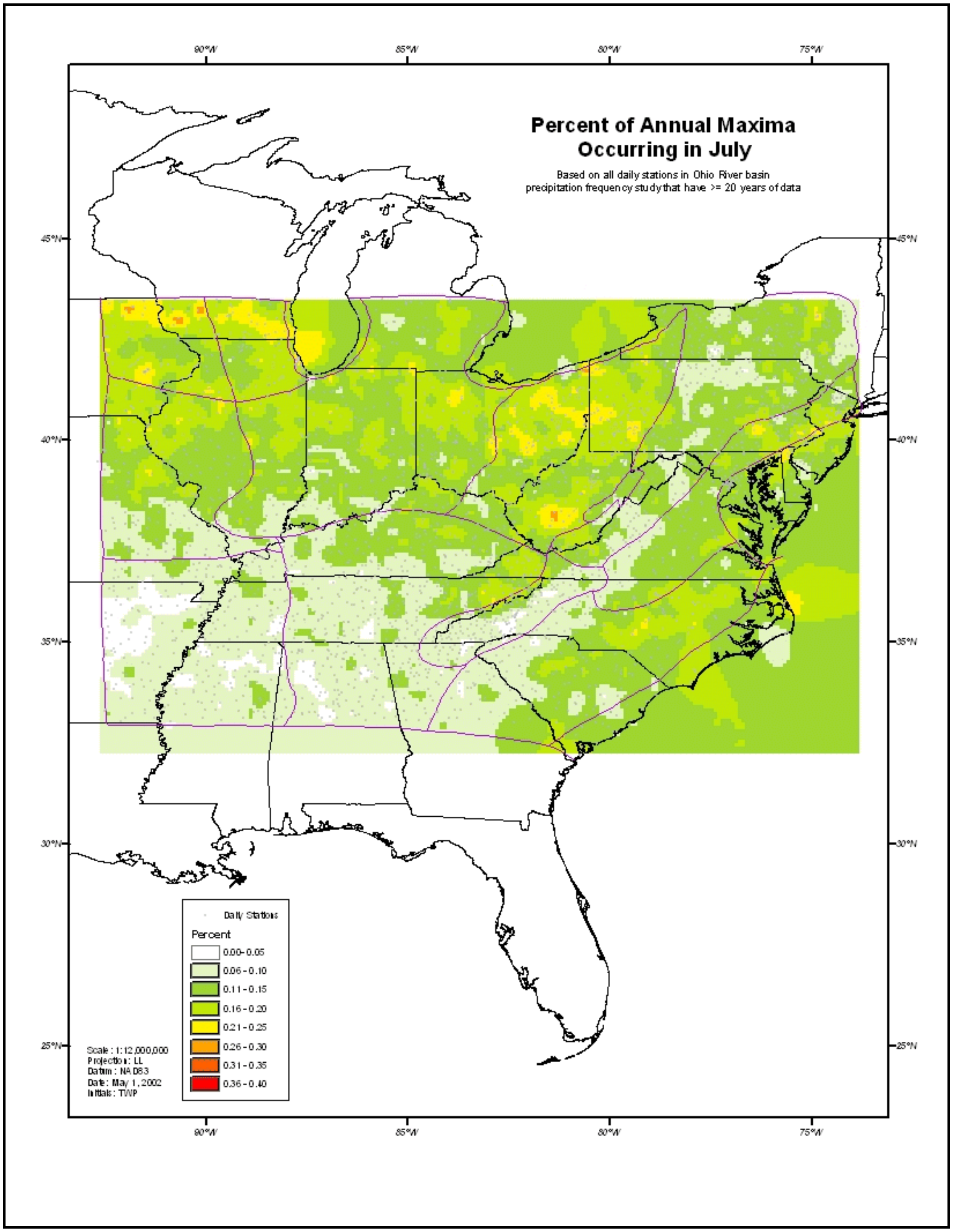
Region	Duration				
	2 hours	3 hours	6 hours	12 hours	**24 hours / 1 day**
1	0.94	1.39	0.43	-1.04	-0.66
2	**2.06**	1.05	0.15	-0.76	-2.30
3	-0.50	-0.07	-0.87	-1.47	-0.08
4	0.52	-1.23	-0.99	-0.60	0.52
5	**2.11**	1.48	1.20	-0.23	**3.05**
6	-0.38	0.67	-0.39	-0.28	-0.95
7	0.43	1.55	**2.69**	**2.02**	**4.90**
8	0.96	1.37	-0.96	-1.03	-1.02
9	0.37	0.47	1.46	0.67	1.68
10	0.59	1.27	0.84	-1.03	1.54
11	1.68	1.49	1.22	1.03	**2.01**
12	-0.09	0.00	0.05	-0.29	0.48
13	-0.81	-0.66	-0.22	0.59	1.20
14	-1.99	-1.88	-0.53	0.56	-1.19
15	0.44	1.15	0.82	0.75	1.89
16	1.56	1.44	**2.67**	**2.77**	**2.42**

4.4 Temporal Distribution

Our methodology for developing temporal distributions of extreme rainfall events has been researched and verified. Our method is based on an Illinois State Water Survey Report (Huff, 1990) and determines the maximum and median precipitation event time distributions for 12, 24 and 72 hour duration events.

Twelve monthly percentage grid maps have been created for the temporal distribution analysis (see Figure 2). The value plotted is the percentage of annual maxima events occurring in a given month for a station. It equals the number of times the annual maxima occurred in that month divided by the number of total annual maxima events occurring for that station, multiplied by 100.

Figure 2. Percent of annual maxima occurring in July



For a given station:
% an. max. events in January = (# of an. max. in January/ # total an. max. events) * 100

For example:
10% of an. max. events occurred in January = (5 an. max events in January/ 50 an. max. events) * 100

Thus for a station with 60 years of record, 60 annual maxima values exist. If 10 annual maxima events occurred in January, and 5 occurred in February then the values plotted would be 16.7% (i.e., (10/60)*100) and 8.3% (i.e., (5/60)*100).

For initial evaluation, these percentages were mapped and spatially interpolated to a grid for each month using a simple inverse-distance-weighting (IDW) technique. Fifteen minute by fifteen minute grid cells were used in order to discern general patterns and eliminate noise in the data.

Some generalized spatial patterns and monthly differences are evident in the percentage grid maps. Time distributions of hourly maximum and median events will be sorted, averaged and plotted by storm area, quartile and duration.

4.5 Precipitation Frequency Data Server

Several changes were made to the Precipitation Frequency Data Server (PFDS) this quarter. A function to extract data for a specific observing site was added. A user can now select an observing site from a pull-down list or by clicking on it from a map. In order to accommodate the observing site selection option, the PFDS interface changed slightly, as did the output page. Other changes include, but are not limited to: output duration changed from 10-days to 60-days, return periods were extended to 1000-years, reference maps were added to the output page to show the surrounding area, the seasonality option was removed, a link to the National Climatic Data Center (NCDC) showing nearby observing sites and sources of climate data was included, and a color-shaded elevation background was added to the state maps.

4.6 Spatial Relations (Depth Area Duration Study)

During the second quarter of 2002, processing of data continued for study areas being used to develop Depth Area Duration (DAD) relationships applicable to basins ranging in area size from 10 to 400 square miles. Currently, 12 study areas are being considered (See Table 2). The areas were selected based on the following criteria: 1) Availability of a dense area of hourly reporting rain gauges; 2) location, as there is a desire to include as many geographically and orographically diverse study areas of the US as possible; and 3) minimum period of record for reporting gauges (at least 15 years of record).

Data has been collected and prepared as shown in Table 2. Also, if additional dense-area-networks are identified, they will be added after the current software development phase of the projected is completed.

Table 2. Dense Area Rain Gauge Networks in D.A.D. Study.

Depth Area Duration Study Areas	**Data Processed**
Walnut Gulch, AZ	✓
Reynolds Creek,ID	✓
Tifton, GA	✓
Hastings, NE	✓
Alamogordo Creek, NM	
Safford, AZ	
Santa Rita, AZ	
Cochocton, OH	✓
Danville, VT	✓
Chicago, IL (NCDC stations)	✓
Riesel, TX	✓

5. Issues

5.1 Seasonality

We reviewed the meaning, utility and computational issues associated with frequency estimates computed by season. Because of the noise in the data, seasonal estimates do not combine to give annual estimates. This incongruity cannot be resolved without arbitrary limits placed on the results. We are uncomfortable with imposing arbitrary limits, but at the same time are concerned about the confusion that may be caused by publishing incongruous results. Furthermore, after checking with a variety of partners, we found no consensus of demand for the estimates or a consensus on how such results would be used. Accordingly we have decided not to prepare seasonal frequency estimates.

5.2 Publication

Printing the final documents is expensive and time consuming. Furthermore, we have found no reasonable way to avoid ongoing infrastructure costs of delivering and billing for the printed documents. Accordingly, we have decided to avoid both the costs and delay by publishing the documents in PDF format on the Internet.

6. Projected Schedule

The following list provides a tentative schedule with completion dates. Brief descriptions of tasks being worked on next quarter are also included in this section.

Data Collection and Quality Control [Complete]
Temporal Distributions of Extreme Rainfall [July 2002]
L-Moment Analysis/Frequency Distribution [August 2002]
Peer Review of Point Estimates [September 2002]
Spatial Interpolation [December 2002]
Precipitation Frequency Maps [January 2003]
Web Publication [January 2003]
Spatial Relations (Depth Area Duration Studies) [January 2003]

6.1 L-Moment Analysis

A comprehensive L-moment statistical analysis will be completed by the next quarter for all datasets through December 2000 for all durations and all regions.

6.2 Trend and Shift Analysis

The dataset will be analyzed for any trends or shifts in annual maxima through time. T-tests will be used to detect any linear trends in annual maxima or in variance, while t-tests, Mann-Whitney tests and Chi-squared tests will be used to determine any shifts in means of annual maxima. After completion of the trend and shift analysis, data quality control will be performed on stations exhibiting a significantly high linear trend and/or shift in the annual maxima time series data.

6.3 Temporal Distributions of Extreme Rainfall

Our methodology for developing temporal distributions of extreme rainfall events has been researched and verified. Our method is based on an Illinois State Water Survey Report (Huff, 1990) and determines the maximum and median precipitation event time distributions for 12, 24 and 72 hour duration events. Time distributions of hourly maximum and median events will be sorted, averaged and plotted by storm area, quartile and duration.

6.4 Spatial Relations (Depth Area Duration Study)

The method to be used for computing the DAD curves has been selected. Software to decode and format the data files and the DAD computations will be developed. If additional dense-area-networks are available, they will be added to the database.

References

Frederick, R.H., V.A. Myers and E.P. Auciello, 1977: Five- to 60-minute precipitation frequency for the eastern and central United States, NOAA Technical Memo. NWS HYDRO-35, Silver Spring, MD, 36 pp.

Hershfield, D.M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years, *Weather Bureau Technical Paper No. 40*, U.S. Weather Bureau. Washington, D.C., 115 pp.

Hosking, J.R.M. and J.R. Wallis, 1997: *Regional frequency analysis, an approach based on L-moments*, Cambridge University Press, 224 pp.

Huff, F. A., 1990: Time Distributions of Heavy Rainstorms in Illinois. Illinois State Water Survey, Champaign, 173, 17pp.

Lin, B. and L.T. Julian, 2001: Trend and shift statistics on annual maximum precipitation in the Ohio River Basin over the last century. Symposium on Precipitation Extremes: Prediction, Impacts, and Responses, 81st AMS annual meeting. Albuquerque, New Mexico.

Miller, J.F., 1964: Two- to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States, *Technical Paper No. 49*, U.S. Weather Bureau and U.S. Department of Agriculture, 29 pp.

Miller, J.F., R.H. Frederick and R.J. Tracy, 1973: Precipitation-frequency atlas of the western United States, *NOAA Atlas 2*, 11 vols., National Weather Service, Silver Spring, MD.